

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2002, 08:25:25 ; Search time 22.96 seconds

(without alignments)  
146.716 Million cell updates/sec

Title: US-09-351-778A-10

Perfect score: 87  
Sequence: 1 MTGSTIAFTDYRNTATGL.....RPPIYRPIGLKPCSLILOYD 87

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	86.2	101	E311_ADE02	P24935 human adeno
2	75	86.2	101	E311_ADE02	O55653 human adeno
3	24	40.2	101	E311_ADE06	P17590 human adeno
4	24	27.6	93	E311_ADE05	P70678 mus musculu
5	6	6.9	92	HRK_MOUSE	002123 poplar mosa
6	6	6.9	121	VNBP_POPPV	0925V0 frankia sp.
7	6	6.9	126	NIFM_PRASE	009320 caenorhabdi
8	6	6.9	136	YR6_CAEEL	P38337 saccharomyc
9	6	6.9	142	YR6_YEAST	P82244 spinacia ol
10	6	6.9	152	RK34_SPTOL	O54792 mus musculu
11	6	6.9	157	HES2_MOUSE	P35429 rattus norv
12	6	6.9	173	HES2_RAT	O9Y543 homo sapien
13	6	6.9	184	YR89_HAELN	P44215 haemophilus
14	6	6.9	194	IM22_HUMAN	O9Y584 homo sapien
15	6	6.9	195	NFTN_MOUSE	P97463 mus musculu
16	6	6.9	248	RIBF_RHOOP	O84990 rhodococcus
17	6	6.9	248	VG12_EBV	P03218 Epstein-Bar
18	6	6.9	263	Y225_METUA	O57678 methanococ
19	6	6.9	280	HES1_HUMAN	O14469 homo sapien
20	6	6.9	281	HES1_RAT	O04666 rattus norv
21	6	6.9	282	HES1_MOUSE	P35428 mus musculu
22	6	6.9	290	HES1_CHICK	O57337 gallus gall
23	6	6.9	290	YGAF_YEAST	P42938 saccharomyc
24	6	6.9	318	FEPB_ECOLI	P14609 escherichia
25	6	6.9	325	YR03_YEAST	P38844 saccharomyc
26	6	6.9	328	IR11_HUMAN	O01638 homo sapien
27	6	6.9	328	SYFA_HELPJ	O92K1F helicobacte
28	6	6.9	338	SYFA_HELPJ	P56146 helicobacte
29	6	6.9	337	HAIR_DROME	P14003 dirosophila
30	6	6.9	350	Y098_SYNY3	O55880 synecocyst
31	6	6.9	361	OXDA_FUSSO	P24552 fusarium so
32	6	6.9	369	SP11_MYXVL	P12393 myxoma viru
33	6	6.9	377	TTL_BOVIN	P38584 bos taurus

34	6	6.9	378	1	HAIR_DROVI	P29303 dirosophila
35	6	6.9	379	1	TTL_PIG	P38160 sus scrofa
36	6	6.9	398	1	HEMT_ECOLI	P09128 escherichia
37	6	6.9	401	1	CSD_PSEPU	O92408 pseudomonas
38	6	6.9	408	1	AATC_CAEEL	O92407 caenorhabdi
39	6	6.9	410	1	NORB_PASMU	O05011 haemophilus
40	6	6.9	411	1	NORB_HAELN	P49219 xenopus lae
41	6	6.9	420	1	MLIC_XENLA	O52655 rhodospirill
42	6	6.9	434	1	BZTC_RHOCA	O52655 rhodospirill
43	6	6.9	435	1	DPK_DROME	O26263 dirosophila
44	6	6.9	435	1	FIXC_RHIME	P09820 rhizobium m
45	6	6.9	440	1	YD0N_HAELN	P45079 haemophilus

## ALIGNMENTS

RESULT	ID	E311_ADE02	STANDARD:	PRT:	101 AA.
AC	P24935:				
DT	01-MAR-1992 (Rel. 21, Created)				
DT	01-MAR-1992 (Rel. 21, Last sequence update)				
DT	01-DEC-1992 (Rel. 24, Last annotation update)				
DE	Early E3A 11.6 kDa glycoprotein.				
OS	Human adenovirus type 2.				
OX	NCBI_TaxID=10515:				
RN	SEQUENCE FROM N.A.				
RP	MEDLINE=81053687; PubMed=6253880;				
RA	Scalia A., Tolleison A.E., Saha S.K., Mold W.S.M.;				
RA	Herisse J., Courtois G., Gallibert F.;				
RT	"Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome.";				
RT	Nucleic Acids Res. 8:2173-2192(1980).				
RN	[2]				
RP	IDENTIFICATION OF PROTEIN.				
RA	MEDLINE=93079877; PubMed=1448922;				
RT	"The E3-11.6K protein of adenovirus is an Asn-glycosylated integral				
RT	membrane protein that localizes to the nuclear membrane.";				
RL	Virology 191:743-753(1992).				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.				
CC	-1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.				
CC	-1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6				
CC	kDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.				
CC	-----				
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sdb.ch">license@sdb.ch</a> ).				
CC	-----				
DR	EMBL: J01917; AAA9222.1.				
FT	Early protein; Glycoprotein; Transmembrane.				
KW	TRANSMEM				
FT	CARBOHYD 41				
FT	62				
FT	POTENTIAL.				
FT	N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	SEQUENCE 101 AA: 11644 MW: 18995 P0921E84B CRC64:				
Query Match	Best Local Similarity	86.2%	Score 75:	DB 1:	Length 101:
Matches	75: Conservative	0:	Mismatches	0:	Indels
DB	1	MTGSTIAFTDYRNTATGLTSLNLPQVAFVNDMSLDMWFSIALMFVCLIMMLIC 60			
DB	1	MTGSTIAFTDYRNTATGLTSLNLPQVAFVNDMSLDMWFSIALMFVCLIMMLIC 60			
OY	61	CLKRRRAPPIYRPI 75			
OY	61	CLKRRRAPPIYRPI 75			
DB	61	CLKRRRAPPIYRPI 75			

```

RESULT 2
E311_ADE06 STANDARD: PRT: 101 AA.
AC 055653:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Early E3A 11.6 kDa glycoprotein.
OS Human adenovirus type 6.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
CC NCBI_TaxID=10534;
RN [1]
RP SEQUENCE FROM N.A.
RA Reischmann H., Scharschmidt E., Geisler B., Hausmann J., Ortmann D.,
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
KDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y16037; CAJ75991.1;
DR Early protein: Glycoprotein; Transmembrane.
KW TRANSMEM 41 62
FT CAROHD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 101 AA; 11613 MW; EF219000939E3B4B CRC64;
SO
Query Match 40.2%; Score 35; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 14e-30;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 MMFSLALMFVCLIMLIMLILCLKRRARRPIYRPI 75
Db 41 MMFSLALMFVCLIMLIMLILCLKRRARRPIYRPI 75
RESULT 3
E311_ADE05 STANDARD: PRT: 93 AA.
AC P17590:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Early E3A 10.5 kDa glycoprotein.
OS Human adenovirus type 5.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
CC NCBI_TaxID=28285;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85092388; PubMed=2981456;
RA Cleaves C., Wold W.S.M.;
RT "DNA sequence of the early E3 transcription unit of adenovirus 5.";
RL Virology 140:28-43(1985).
RN [2]
RP COMPLETE GENOME.
RA MEDLINE=92087470; PubMed=1727603;
RA Chroboczek J., Bieber F., Jacot B.;
RT "The sequence of the genome of adenovirus type 5 and its comparison
RT with the genome of adenovirus type 2.";
RL Virology 186:280-285(1992).
CC -1- SUBCELLULAR LOCATION: INTEGRAL NUCLEAR MEMBRANE PROTEIN.
CC -1- PTM: N-GLYCOSYLATED AND PROBABLY ALSO O-GLYCOSYLATED.

```

```

CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER AD2 E3A 11.6
CC kDa, AD5 E3A 10.5 kDa, AD3 E3 9 kDa, AND AD7 E3 7.7 kDa.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M73260; NOT_ANNOTATED_CDS.
DR EMBL: X03002; CA26784.1;
DR PIR: A05245; ERAD53.
KW Early protein: Glycoprotein; Transmembrane.
FT TRANSMEM 34 55
FT CAROHD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 93 AA; 10523 MW; 008AD087AAB45ABF CRC64;
SO
Query Match 27.6%; Score 24; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 8.7e-19;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 MMFSLALMFVCLIMLIMLILCLKRRARRPIYRPI 64
Db 34 MMFSLALMFVCLIMLIMLILCLKRRARRPIYRPI 57
RESULT 4
HRK_MOUSE STANDARD: PRT: 92 AA.
AC P70678:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Activator of apoptosis harakiri (neuronal death protein dp5).
GN HRK OR DP5 OR BID3.
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CC NCBI_TaxID=10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA Imaizumi K.;
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ACTIVATES APOPTOSIS AND INTERACTS SELECTIVELY WITH
CC SURVIVAL-PROMOTING PROTEINS BCL-2 AND BCL-XL (BY SIMILARITY).
CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR THE INDUCTION OF CELL
CC DEATH (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D83697; BAA12065.1;
DR EMBL: D83698; BAA12066.1;
DR MGD: MGI:1201608; B1d3.
DR InterPro: IPR000712; Bcl_2.
DR PROSITE: PS01259; BH3; 1.
RN APOPTOSIS.
FT DOMAIN 33 47 BH3.
FT SEQUENCE 92 AA; 10078 MW; 95760B9EBA88FIDE CRC64;
SO
Query Match 6.9%; Score 6; DB 1; Length 92;

```

Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 RRRARP 69  
| | | | |  
DB 50 RRRARP 55

## RESULT 5

VNBP\_POPMV STANDARD; PRT; 121 AA.  
ID VNBP\_POPMV  
AC 002123;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DE 01-OCT-1993 (Rel. 27, Last annotation update)  
DE 14 kDa protein (Putative nucleic acid-binding protein).  
OS Poplar mosaic virus (isolate ATCC PV275) (PMV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlavirus.  
NCBI\_TaxID=31709;  
RN 11)  
RP SEQUENCE FROM N.A.  
RX MEDLINE-92333281; PubMed-1629709;  
RA Henderson J., Gibbs M.J., Edwards M.L., Clarke V.A., Gardner K.A.,  
RA Cooper J.I.;  
RT "Partial nucleotide sequence of poplar mosaic virus RNA confirms its  
RT classification as a Carlavirus.";  
RL J. Gen. Virol. 73:1887-1890(1992).  
CC -1- FUNCTION: MAY ACT AS A REGULATORY FACTOR DURING VIRAL  
CC TRANSCRIPTION.  
CC -1- SIMILARITY: TO NUCLEIC ACID-BINDING PROTEINS FROM OTHER  
CC CARLAVIRUSES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X65102; CAA46227.1; -;  
CC EMBL: D13364; BAA02629.1; -;  
CC PIR: J01646; J01646.  
CC PIR: S23876; S23876.  
CC InterPro: IPR002568; Carla\_C4.  
CC Pfam: PF01623; Carla\_C4; 1.  
CC ZINC-finger; DNA-binding.  
CC FT ZN-FING 62 83  
CC FT C4-TYPE (POTENTIAL).  
CC SO SEQUENCE 121 AA; 14451 MW; 705205527BF7D3CF CRC64;

Query Match 6.9%; Score 6; DB 1; Length 121;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 KRRRAR 68  
| | | | |  
DB 52 KRRRAR 57

## RESULT 6

NIFM\_FRASE STANDARD; PRT; 126 AA.  
ID NIFM\_FRASE  
AC 092570;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Nitrogenase stabilizing/protective protein nifm.  
GN NIFM.  
OS Frankia sp. (strain EuIK1).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Frankineae; Frankiaceae; Frankia.  
NCBI\_TaxID=47227;

RN 11)  
RP SEQUENCE FROM N.A.  
RC STRAIN-EuIK1;  
RA Chung-Sun A., Ji-Tee K., Won-Jin K., Won-Young Y.;  
RT "Nif-gene organization and nucleotide sequences from Frankia EuIK1  
RT strain.";  
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative  
CC damage (by similarity).  
CC -1- SUBUNIT: Homotrimer; associates with nifD (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF119361; AAD17267.1; -;  
CC DR Nitrogen fixation.  
KM SEQUENCE 126 AA; 14287 MW; 295D8C9102FD317E CRC64;  
SO SEQUENCE

Query Match 6.9%; Score 6; DB 1; Length 126;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TTATGL 20  
| | | | |  
DB 80 TTATGL 85

## RESULT 7

YRN6\_CAEEL STANDARD; PRT; 136 AA.  
ID YRN6\_CAEEL  
AC 009420;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Hypothetical 14.2 kDa protein R0781.6 in chromosome X.  
GN R0781.6  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN 11)  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Kerhaw J.;  
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z48621; CAA8543.1; -;  
CC WormPep: R0781.6; CE01632.  
CC InterPro: IPR002601; C6.  
CC Pfam: PF01681; C6; 1.  
CC Hypothetical protein.  
CC SO SEQUENCE 136 AA; 14169 MW; D7F47A83F3F81787 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 136;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 T1APT 10  
111111  
DB 31 T1APT 36

## RESULT 8

YB9E\_YEAST STANDARD; PRT; 142 AA.  
AC P38337;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DE 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Hypothetical 16.4 kDa protein in POP4-SHM1 intergenic region.  
GN YBR258C OR YBR1726.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RX MEDLINE=9320397; PubMed=8465606;  
RA Dolnon F., Bileau N., Crouzet M., Aigle M.;  
RT "The complete sequence of a 19,482 bp segment located on the right  
arm of chromosome II from Saccharomyces cerevisiae.";  
RL Yeast 9:189-199(1993).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X70529; CAA49922.1; -;  
DR EMBL; Z66127; CA885221.1; -;  
DR PIR: S32960; S32960.  
DR SCD: S0000462; YBR258C.  
KW Hypothetical protein.  
SQ SEQUENCE 142 AA; 16429 MW; 12E73F0C3DD25D7 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 142;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GLTSL 24  
111111  
DB 71 GLTSL 76

## RESULT 9

RK34\_SPTOL STANDARD; PRT; 152 AA.  
AC P82244;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 50S ribosomal protein L34, chloroplast precursor.  
GN RPL34.  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;  
RN [1]  
RP SEQUENCE FROM N.A.; SEQUENCE OF 92-101, AND MASS SPECTROMETRY.  
RC STRAIN-CV. ALVARO; TISSUE=leaf;  
RX MEDLINE=20435798; PubMed=10874046;  
RA Yamauchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the 50 S subunit of an organelle ribosome (chloroplast).";

RL J. Biol. Chem. 275:28466-28482(2000).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -1- SUBCELLULAR LOCATION: Chloroplast.  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- MASS SPECTROMETRY: MW=67.1; METHOD=Electrospray; RANGE=92-152.  
CC -1- SIMILARITY: BELONGS TO THE L34P FAMILY OF RIBOSOMAL PROTEINS.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: AF238221; AAF6457.1; -;  
DR InterPro: IPR000271; Ribosomal\_L34.  
DR Pfam: PF00468; Ribosomal\_L34; 1.  
DR PROSITE: PS00784; RIBOSOMAL\_L34; FALSE NEG.  
KW Ribosomal protein; Chloroplast; Transit peptide; rRNA-binding.  
FT CHAIN 92 152 50S RIBOSOMAL PROTEIN L34.  
SQ SEQUENCE 152 AA; 16095 MW; 755A990D441ADB18 CRC64.

Query Match 6.9%; Score 6; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 LKRRRA 67  
111111  
DB 127 LKRRRA 132

## RESULT 10

HES2\_MOUSE STANDARD; PRT; 157 AA.  
ID HES2\_MOUSE  
AC O54792;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Transcription factor HES-2 (Hairy and enhancer of split 2).  
GN HES2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/J;  
RX MEDLINE=98234545; PubMed=9570950;  
RA Nishimura M., Isaka F., Ishibashi M., Tomita K., Tsuda H.,  
RA Nakaniishi S., Kageyama R.;  
RT "Structure, chromosomal locus, and promoter of mouse Hes2 gene, a  
homologue of Drosophila hairy and enhancer of split.";  
RL Genomics 49:69-75(1998).  
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH  
PROTEIN FOR THEIR TRANSCRIPTION.  
CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX  
WITH A CO-REPRESSOR PROTEIN (GROUCHO).  
CC -1- SUBCELLULAR LOCATION: Nucleus.  
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A  
HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNG),  
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).  
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL  
REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A  
TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY  
HAIR-RELATED PROTEINS.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL: AB009967; BAA24091.1; -  
CC MGD: MGI:1098624; Hes2.  
CC InterPro: IPR003015; HLH\_MYC.  
CC InterPro: IPR001092; HLH\_dlm.  
CC InterPro: IPR003650; Orange.  
CC Pfam: PF00010; HLH; 1.  
CC SMART: SM00353; HLH; 1.  
CC DR SMART: SM00511; ORANGE; 1.  
CC DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; 1.  
CC DR Transcription regulation; DNA-binding; Nuclear protein; Repressor.  
CC FT DNA\_BIND 13 26 BASIC DOMAIN.  
CC FT DOMAIN 139 148 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
CC FT DOMAIN 154 157 WRW MOTIF (REQUIRED FOR ACTIVITY)  
CC (BY SIMILARITY).  
CC FT SEQUENCE 157 AA; 17231 MW; 570A0C67F4992EA7 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 157;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 KRRRAR 68  
| | | | |  
DB 23 KRRRAR 28

RESULT 11  
HES2\_RAT STANDARD; PRT: 157 AA.  
ID HES2\_RAT  
AC P33429;  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Transcription factor HES-2 (Hairy and enhancer of split 2).  
GN HES2 OR HES-2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryonic brain;  
RX MEDLINE=9335886; PubMed=8354270;  
RA Ishibashi M., Sasai Y., Nakanishi S., Kageyama R.;  
RT Molecular characterization of HES-2, a mammalian helix-loop-helix  
RT factor structurally related to Drosophila hairy and Enhancer of  
RT split.  
RL Eur. J. Biochem. 215:645-652(1993).  
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH  
CC PROTEIN FOR THEIR TRANSCRIPTION.  
CC -1- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX  
CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A  
CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACCNG).  
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).  
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRW MOTIF IS A TRANSCRIPTIONAL  
CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A  
CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY  
CC HAIRY-RELATED PROTEINS.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL: D14029; BAA03118.1; -  
CC PIR: S35037; S35037.  
CC TRANSFAC: T01650; -  
CC InterPro: IPR003015; HLH\_MYC.  
CC InterPro: IPR001092; HLH\_dlm.  
CC InterPro: IPR003650; Orange.  
CC Pfam: PF00010; HLH; 1.  
CC SMART: SM00353; HLH; 1.  
CC DR SMART: SM00511; ORANGE; 1.  
CC DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; 1.  
CC DR Transcription regulation; DNA-binding; Nuclear protein; Repressor.  
CC FT DNA\_BIND 13 26 BASIC DOMAIN.  
CC FT DOMAIN 139 148 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
CC FT DOMAIN 154 157 WRW MOTIF (REQUIRED FOR ACTIVITY)  
CC (BY SIMILARITY).  
CC FT SEQUENCE 157 AA; 17028 MW; B5D621E814AE0369 CRC64;

Query Match 6.9%; Score 6; DB 1; Length 157;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 KRRRAR 68  
| | | | |  
DB 23 KRRRAR 28

RESULT 12  
HES2\_HUMAN STANDARD; PRT: 173 AA.  
ID HES2\_HUMAN  
AC Q9Y543; Q9Y542;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Transcription factor HES-2 (Hairy and enhancer of split 2).  
GN HES2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howden P.;  
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF GENES THAT REQUIRE A BHLH  
CC PROTEIN FOR THEIR TRANSCRIPTION (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A  
CC HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACCNG).  
CC RATHER THAN THE CANONICAL E-BOX (CANNTG).  
CC -1- DOMAIN: THE CARBOXYL-TERMINAL WRW MOTIF IS A TRANSCRIPTIONAL  
CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A  
CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY  
CC HAIRY-RELATED PROTEINS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurturin precursor.
GN NRTN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 96-110; 127-135; 155-177 AND
RP 181-190.
RX MEDLINE=971100947; PubMed=8945474;
RA Kotzbauer P.T., Lampe P.A., Heuckeroth R.O., Golden J.P.,
RA Crendon D.J., Johnson E.M. Jr., Milbrandt J.;
RT "Neurturin, a relative of glial-cell-line-derived neurotrophic
RT factor ";
RL Nature 384:467-470(1996).
CC -1- FUNCTION: SUPPORTS THE SURVIVAL OF SYMPATHETIC NEURONS IN CULTURE.
CC MAY REGULATE THE DEVELOPMENT AND MAINTENANCE OF THE CNS. MIGHT
CC CONTROL THE SIZE OF NON-NEURONAL CELL POPULATION SUCH AS
CC HAEMOPOIETIC CELLS.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: WIDESPREAD DISTRIBUTION.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY. GDNF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U78109; AAC52954.1; -.
DR HSSP: Q07731; IAGQ.
DR MGD: MGI:108417; Nrtu.
DR InterPro: IPR002400; GF_cysknol.
DR InterPro: IPR001839; TGF-beta.
DR Pfam: PF00019; TGF-beta; 1.
DR PRINTS: PRO0438; GFCYSKNOT.
DR SMART: SM00204; TGFBR; 1.
DR PROSITE: PS00250; TGF-BETA_1; FALSE_NEG.
KW Growth factor; Signal.
FT SIGNAL 1 19
FT PROPEP 20 95
FT CHAIN 96 195
FT DISULFID 101 163
FT DISULFID 128 192
FT DISULFID 132 194
FT DISULFID 162 162
SQ SEQUENCE 195 AA; 22219 MW; ABE21BB35D417448 CRC64;

```

```

Query Match 6.9%; Score 6; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 64 RRRARP 69
DB 91 RRRARP 96

```

Search completed: June 21, 2002, 08:25:25  
Job time: 310 sec